

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)**SciVerse ScienceDirect**

Procedia Computer Science 10 (2012) 282 – 291

**Procedia**  
Computer Science

The 3<sup>rd</sup> International Conference on Ambient Systems, Networks and Technologies  
(ANT)

## MobilePaperAccess: Ubiquitous Paper-based Interfaces for Mobile Interactions

Yun ZHOU, Bertrand DAVID\*, René CHALON

*Université de Lyon, CNRS, Ecole Centrale de Lyon, LIRIS, UMR5205,  
36 avenue Guy de Collongue, F-69134 Ecully Cedex, France  
E-mail [Bertrand.David@ec-lyon.fr](mailto:Bertrand.David@ec-lyon.fr), Tel. (33)472186581, Fax (33)472186443*

---

### Abstract

MobilePaperAccess is a wearable camera-glasses system with tangible user interface allowing mobile interaction. We access the digital information from a paper interface extending the input space. The system is devised to validate our concepts of Environment Dependent Interface (EDI) and Environment Independent Interface (EII), which focus on enabling people to access their personal data as well as public resources at any time and in any place. In this paper, we propose a continuum from physical interface to digital interface in relation with EDI and EII, and we present design, implementation and evaluation aspects of our MobilePaperAccess system. We compare two interfaces (EDI and EII) and three input techniques (finger input, mask input and page input).

© 2011 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of [name organizer]  
Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

Keywords: Wearable computing, input techniques, user interface, finger tracking, ubiquitous computing, paper interaction, tangible user interface, mobility

---

### 1. Introduction

With the development of mobile sensors and small computing devices, the availability of innovative input techniques and interfaces has greatly increased. These new interaction methods are changing people's life by facilitating their tasks in daily life and in working places, and enabling them to access personal data as well as public resources at any time and in any place. In order for mobile users to interact with their environment, we propose three types of interfaces [1]: In-environment Interface (IEI), Environment Dependent Interface (EDI) and Environment Independent Interface (EII). Figure 1 shows

the relations between the interfaces and three main elements: Environment, User and Devices, as well as the contextualization provided by these interfaces. With the In-Environment Interface (IEI), the user is in the nomadic state, i.e., without any personal IT device. The environment provides all the interaction support for input and output. In this situation, a fix webcam and a wall video projector are appropriately located to allow in-environment interaction. The user uses his/her hands to interact with the public information available from a public wall, e.g., searching and browsing. The context awareness is determined by the environment (physical location) and the application used (public transportation information). The other two user interfaces (EDI and EII) rely on the user's wearable computer. The Environment Dependent Interface (EDI) refers to the strong relationship between the interface and the in-environment information, whereas the Environment Independent Interface (EII) refers to the relation between the interface and the personal information that can be used for contextualization.

In the smart city [2], the IEI, EDI and EII can solve the identical as well as distinct problems met by the user. For example, the nomadic user (without wearable computer) can use the In-Environment Interface of the bus shelter (figure 2b) to ask for the schedule of the next bus. If he is equipped with the EDI interface (figure 2c) he can use the wall of the bus shelter to contextualize his location and receive the next bus's schedule in his goggle attached to the small screen, or if he is equipped with the EII interface (figure 2a), he can contextualize and indicate his location with the projected city map via pico-projector.

In this paper, we present MobilePaperAccess, a ubiquitous paper-based interface for mobile interactions based on the EDI and EII concepts. We employ the following wearable configuration: a small screen attached to a goggle to provide visual information, a webcam to pick the input signal, and a laptop as the calculating device. Our objective is to create a true contextualization based on the user's location or independent of it, which is more effective and adaptive to users' information needs by taking advantage of dynamic and physical environmental characteristics.

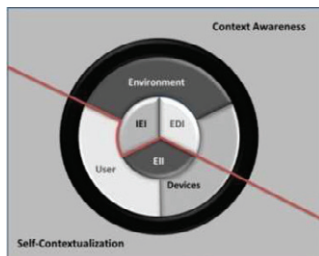


Fig. 1. The Principal relationships between Environment, User and Devices.

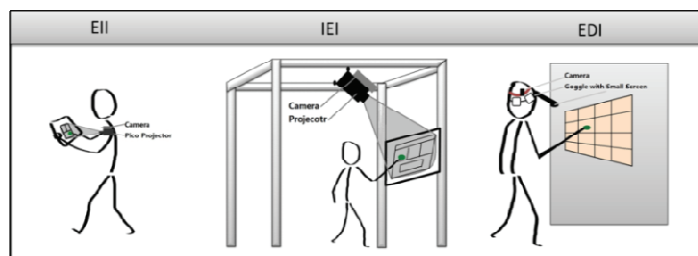


Fig. 2. (a) EII; (b) IEI; (c) EDI.

## 2. Related work

The ubiquitous computing covers a large number of aspects. We only address wearable input and output techniques and paper interaction in this paper.

### 2.1. Wearable Input and Output Techniques

The miniaturization of mobile and wearable devices makes the ubiquitous computing possible, searching for the wearable input modality and mobile output becomes the point of research. The technology of Skinput [3] employs the user's body as the interactive surface such as the touch pad with the bio-acoustic sensors and projector, which provides an always-available interface [4]. Minput [5] offers an input technique via gestures like flicking and twisting, which is realized by two optical tracking

sensors on the back of a small device. According to the survey of pico-projectors for pervasive computing [7], four ways of interacting with a personal projector system have been identified: input on the projector, movement of the projector, direct interaction with the projection and manipulation of the projection surface. MotionBeam [6] is a novel interaction metaphor, based on the input via the projector movement: the user can navigate by changing the location and orientation of the projector. Unlike the movement interaction of projector, many researchers focus on manipulating the dynamic projection surface. OmniTouch [8] allows bare hand gestures as input, and SixSense [9] explores and proposes the marked fingers gestures as input. Both use the dynamic projection interface.

We focus on the visual output for feedback. With the advent of the pico-projectors, the researchers have investigated the wearable camera-projector system in recent years, OmniTouch [8], SixthSense [9] and Brainy hand [12] are all based on this system and use the projection approach. However, the projective information both on palm and on the wall is easy to be seen by others. As one of the wearable output visual displays, head worn displays [10][11] have been used to present the feedback. With these micro optical output devices, other people only see the transparent glasses without accessing the digital information. The user privacy is thus protected in public areas.

## 2.2. Paper Interaction

Paper interaction is one of the tangible user interfaces [13]. The studies on the paper interaction and paper interfaces [14][15][16] focused on the augmented reality [17], making an effort to merge the use of the paper with the digital information and data. The researchers mark the paper with special markers, and then use the camera to recognize and detect both the motion of paper and other input techniques. Paper Windows [14] describes a projecting window prototype that can simulate the manipulation of the digital paper displays. This system takes the paper motion and finger pointing gestures as the input. The user can thus perform tasks by interacting with paper documents using his fingers, hands and stylus. Quickies [15] is the augmenting sticky notes as an I/O interface. The DisplayObjects [16] proposes a workbench that allows the user to interact with projected information on the physical object. Whereas these studies all investigated -the large display interaction or the desktop interaction, we focus on the paper interaction in the mobile situation.

## 3. Continuum for EDI and EII

We propose a continuum (Figure 3) that spans from physical interface to digital interface. The physical interface surface is static and inflexible in the form of unitary planar, regarded as the uniplanar. The physical-digital interface incorporates the physical interface with the digital one, for which the paper interface (physical interface) has been augmented with the projected interface (digital interface). The latter is based on the multiplanar, in the form of which the interactive elements are organized logically in dynamic multilayer windows. In the EDI system, the interface relies deeply on the environment and the context information, such as the location. Based on this dependence, the EDI builds on the physical interface and the physical-digital one. In the EII system, the interface spans from the physical interface to the digital one.

## 4. MobilePaperAccess

**Input Techniques:** We propose three input techniques as the figure 4 shows: finger input, mask input and page input. Our finger input technique is to let the user's finger hover for a second; the selection signal can be generated via a span. When the user points at a button, he/she needs to remain within the position of the button for a time interval. We propose the hover gesture with a three-state input model

[18] (Figure 5). In the zero state, the finger is beyond the reach of webcam's vision; movements of the finger have no effect on the system. As the finger enters the range of the webcam, the system starts to track it and the tracking symbol is the tip point of user's index finger. The two actions "Hovering for a Second" and "Stop Hovering" are closely knit; the "Stop Hovering" is non-substitutable and strongly combined with the preceding action. Thus, the return path from state 2 to state 1 is drawn in gray. For the mask input, we propose a mask consisting of a green frame and a wand, which shares the same hovering method with the finger input. The real information on the paper can be read in the center of the mask. For the page input, we use a booklet of several pages where each page has an AR-toolkit tag [19]. The index in front of the booklet is for the user easier access to appropriate page.

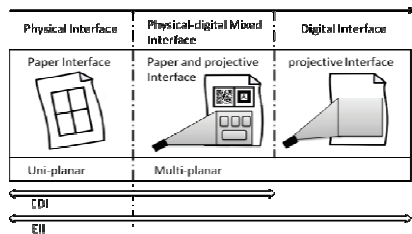


Fig. 3. The continuum

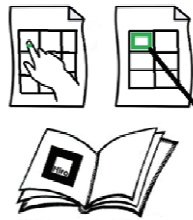


Fig. 4. Three inputs

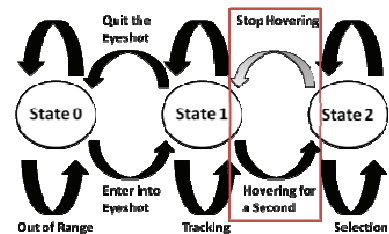


Fig. 5. Three states of hover gesture

**Paper Surface:** According to the human factors (Figure 6), the angle of eye rolling is  $15^\circ$  of comfort and  $35^\circ$  of maximum horizontally and  $30^\circ$  up and  $35^\circ$  down vertically. The average forward grip reach is 74.3cm [20]. The interactive surface held in hand should be less than  $34.64\text{ cm} \times 16.08\text{ cm}$  in size when the reading distance is 30cm. Thus, we choose an A4 ( $29.7\text{ cm} \times 21.0\text{ cm}$ ) paper as interface, and organize the layout under the comfortable range.

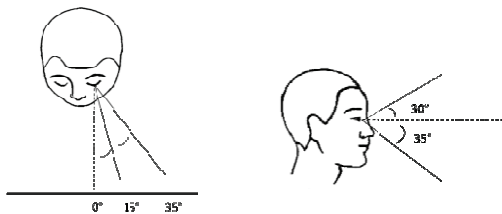


Fig. 6. The angle of eye rolling horizontally and vertically



Fig. 7. (a) Wearable configuration (b) Goggle with small screen

## 5. Implementation

The perspective of our MobilePaperAccess includes the paper interactive surfaces augmented with the color markers, a colored sticker located on user's index finger, AR-toolkit tags, the webcam to capture the motion of the marked index finger or capture the AR-toolkit tags, the goggle with small screen to present the digital information, and a laptop.

**Augmented Paper:** Unlike the devices where the input takes place directly on the display surface, our display and the input are separated. Each paper interactive surface is either augmented with the color markers or with AR-toolkit tags. Two color markers in a diagonal position (Figure 8) shape a rectangle, which can be tracked by the webcam. As long as the webcam recognizes this rectangle shape, the grid within the shape is considered as icons and can be activated by pointing. The user is unaffected even if he rotates or slightly moves the paper during interaction. The booklet for interaction in EII is augmented with AR-toolkit tags and each page has a tag as the identity.

**Goggle with Small Screen:** The feedbacks are presented on the small screen fixed on the right side or left side of the goggle. With the limitation of the screen size, the display area is divided into two parts: the main display area and the auxiliary area (figure 9). The main display area displays the information completely, while the auxiliary one displays the brief response of the information in the form of keyword or tips, permitting a quick and just-in-time understanding of the user.

**Motion of Finger and Mask:** We fix a unique color marker on the index finger of the user, which can be tracked by the CamShift algorithm in real-time. As shown in Figure 10, we record the trace of the color marker and count the number of points in one area such as the grey zone. If the number meets our predefined condition, we regard this action as a pointing. For mask input, we calculate the central point of the mask as the tracking point, which is counted in the same way of finger input.

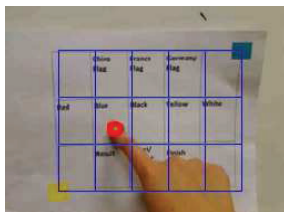


Fig. 8. Colored augmented paper

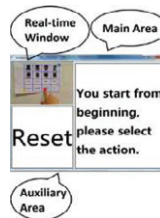


Fig. 9. Visual feedback in small display

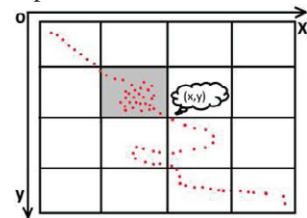


Fig. 10. Motion of the index finger

**Research Team Management Application (RTMA):** To prove the concepts of EDI and EII we developed an application with the objective of managing research team members' exchanges. The scenario with EDI is the following. A research team member wants to consult another member. When he/she arrives at the lab, he/she finds that this person is out. He/she then walks close to the door and starts to use RTMA. He interacts with the paper on the door, looks for an appropriate time and selects it to check the schedule. After marking the decision, he asks for the appointment and gets a feedback from the system. For the same scenario with EII, the user takes out a customized paper or a booklet and holds it in his hand to interact.

## 6. Evaluation

In order to obtain a deeper understanding of our input modalities, we designed a structured evaluation comparing our three input techniques (finger input, mask input and page input) and two interfaces (EDI and EII). We formed four cases as shown in table 1. For cases A and B, the participants stood, whereas for cases C and D, the users sit or stood in a free way to simulate the mobility.

**Participants:** 12 student participants were recruited, including 7 males and 5 females, aged between 19 and 29 (mean of 23.2), and 157cm to 188cm tall (average of 171.8 cm). All participants had the experience of using the mobile devices, 6 of them had the knowledge on HCI. All of them except one were right-handed.

**Procedure:** We provided two types of programs (Toy Application and True Application), each of which contained four cases. Thus, each participant had to finish  $2 \times 4 = 8$  cases. The toy application was a flag game for practicing, by which the participants could choose the flag of interest, then choose its color composition, and finally check the results. The goal of introducing the toy application was to help the participants to get used to the interaction techniques and interfaces. They could play the flag game several times until they felt competent at the eventual true tasks. The evaluation began with an explanation of the protocol in a text form. The questionnaire attached to the protocol contained two parts: the first covered the background information about their familiarity with mobile devices and HCI, as well as basic personal details, that the users had to answer before the test; the second part provided questions in Likert scale form [21], that the users had to finish during and after the test. Next, after practicing several times with the toy application, the users started the true test of RTMA. We employed a within-subjects design, and the order of the cases was counterbalanced with a  $4 \times 4$  Latin Square [22]. All the participants performed the tasks. Each participant was asked to complete two tasks once. They were instructed to check two different researchers' schedules and ask for a meeting with them as quickly and accurately as possible.

Table 1. Four test cases (A, B, C, D)

Interfaces	Input Techniques		
	Finger	Mask	Book/Page
EDI	√ Case A	√ Case B	Not Studied
EII	√ Case C	Not Studied	√ Case D

**Variables:** The independent variables are input techniques (finger input, mask input and page input) and the interfaces (EDI and EII). The dependent variables are interaction time. The interaction time starts from user's correct interaction to his correctly stopping each task. We drew the access time and regarded it as the span from starting the application to user's first interactive action. We also recorded all the errors and found out the reason of the interaction error. The applications recorded the user's input automatically.

## 7. Main Results

**Interaction Time:** In order to know whether any significant difference among input techniques and interfaces exists, we used the Mann Whitney U test [23]. There were no statistically significant differences ( $p>0.05$ ) between the finger input and mask input under the same EDI case and between the EII and EDI with the same finger input. In contrast, significant differences ( $p<0.05$ ) existed between the finger input and book input under the same EII case. The interaction time of page input with EII took a clearly longer time than other cases (Figure 11).

**Access Time:** Figure 12 shows the average time of the access time in each trial. There is no significant difference between cases A, B and C; which only took less than 8 seconds to access. In turn, the access time of case D is nearly two times longer than the other cases.

**Interaction Errors:** Through observation, we found that the reasons of the errors are mainly due to the locomotion of users, the misunderstandings of tasks and attempts beyond the tasks. These three errors were counted respectively. Among these, the locomotion error is the interaction error. The number of interaction errors with EDI is less than with EII, and less with finger input than with page input (Fig. 13).

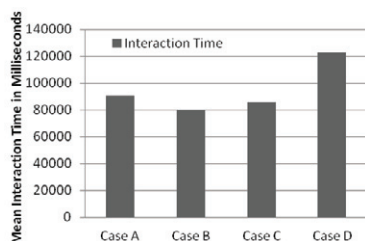


Fig. 11. Mean interaction time.

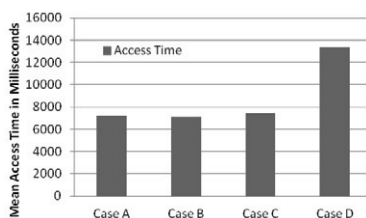


Fig. 12. Mean access time.

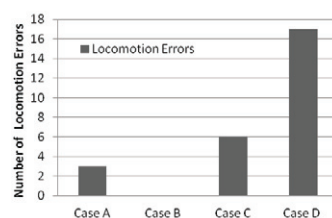


Fig. 13. Locomotion errors.

**User Satisfaction:** We had five levels (1-Strongly disagree, 2-Disagree, 3-Neither agree nor disagree, 4-Agree, 5-Strongly agree) of the Likert items to describe the easiness of learning and the convenience of interaction. Table 2 gives the average scores of each case. All the participants thought that it was not hard to learn and perform (the mean scores are all above 3) and that the interaction in case B was the easiest to learn and the most convenient to perform.

**User Comments:** For case A, i.e. the finger input with the EDI, four participants tired lifting the arm after operating for a while, which lead them to interact unsteadily with finger. Two participants said that the fixed position was efficient and convenient for the interaction. Besides, two participants commented on a physical chain reaction effect when moving the arm and finger resulted in a tiny movement of the camera fixed on the head. For case B, one said that the frame of the mask made it easy to choose and select items, while another user could not work well with the frame's angle of the marker. Two people got their arm tired. For case C, more than half of the participants commented on the long duration of lifting the arm and unsteady finger. They thought that it was not easy to hold the interface in hand steadily. Two participants experienced the chain reaction effect. For case D, four participants found that the search for the right page to interact felt less easy when the pages of the booklet increased, and that returning to the index each time was not convenient. Only one participant mentioned feeling the chain reaction. One user

preferred the marker interaction for a faster and more sensitive interactive experience. For the devices, six participants felt the screen was small to read, which made them somewhat tired.

Table 2. Mean score of user satisfaction of toy application and true application in the four cases

	Toy Application	True Application
Case A	4	4.1
Case B	4.4	4.4
Case C	4	3.5
Case D	3.5	3.5



**Discussions:** Case B has the best performance overall, with the shortest interaction time, the shortest access time, no locomotion error and the best satisfaction. Compared with case A, B has fewer participants expressing arm tiring because the band with mask is more comfortable than the lifted hand. Case A performs better than case C; they almost have the same interaction time and access time except that A has a better score of satisfaction and the A had fewer participants reporting arm tiring because of the fixed and stable interface. In turn, case C performs better than case D for its shorter interaction time, shorter access time, less interaction errors and better satisfaction score. Case D is most influenced by locomotion errors overall. We found that the more the pages, the harder the selection action is. Because of searching pages via returning to the index, the input technique in case D leads the users to an unsteady interaction state.

In order to decrease the locomotion errors in the system with EDI and EII, we propose two solutions. One is to decrease the paper size and increase the paper hardness. Expanded paper size can avoid the problem of fat finger. However, it is easy to let part of the paper out of the field of webcam. In addition, the holding strength can result in the bending of the paper, which would decrease the recognition of webcam and bring about the same interaction problem as the locomotion errors. However, an improved hardness of the paper can compensate for this effect. We can choose the cardboard as the paper interactive surface of the EII. The physical paper interface has low multiplexed ability; the selected items are physical and can't be changed dynamically. If we reduce the space and size of paper interface, the contents also decrease. In order to balance the size and the content, we propose to introduce the AR-toolkit tags into the paper interface to provide the EDI with a physical and dynamical interface. We also found that raising hands at eye level was tiring after a period of time and that the chain reaction decreased the efficiency of interaction. Thus, we propose to change the position of the webcam from forehead to the chest for lowering the elevation of hands and provide the stable experience.

## 8. Conclusion and Future Work

In this paper, we present our taxonomy of mobile user interfaces and main concepts of Environment Dependent Interface and Environment Independent Interface. We proposed, designed and implemented MobilePaperAccess User Interface based on a webcam, a small screen attached to a goggle and a laptop as calculating device. Users can interact with the paper using fingers, masks and pages. This interaction modality supports the privacy and avoids the problem of fat fingers and blocking of focus point by shadow and finger itself. The quantitative and qualitative results show that the main interaction error is the locomotion error and the mask input with the EDI has the best performance. For future work, we plan to investigate the physical and digital combined paper interface with AR-toolkit tags via the devices of pico-projector and webcam. Furthermore, additional input techniques of hand gestures such as the pinch gesture will be studied.

## References

- [1] Zhou Y, David B, Chalon R. Innovative user interfaces for wearable computers in real augmented environment. In: Jacko JA, editor. *HCI International 2011*, Berlin Heidelberg: Springer-Verlag; 2011, p. 500-509. ISBN 978-3-642-21604-6.
- [2] David B, Zhou Y, Xu T, Chalon R. Mobile User Interfaces and their Utilization in a Smart City. *The 2011 International Conference on Internet Computing (ICOMP'11) as part of WorldComp'2011 Conference*, CSREA Press; 2011, p. 383-388.
- [3] Harrison C, Tan D, Morris D. Skinput: appropriating the body as an input surface. *In Proceedings of the 28th international conference on Human factors in computing systems*, 2010, p. 453-462.
- [4] Morris D, Tan D, Saponas TS. Emerging Input Technologies for Always-Available Mobile Interaction. *Foundations and Trends in Human-Computer Interaction*; 4: 4, 2010, p. 245-316.

- [5] Harrison C, Hudson S. E. Minput: enabling interaction on small mobile devices with high-precision, low-cost, multipoint optical tracking. *In Proceedings of the 28th international conference on Human factors in computing systems*, 2010, p. 1661–1664.
- [6] Willis KDD, Poupyrev I, Shiratori T. Motionbeam: a metaphor for character interaction with handheld projectors. *In Proceedings of the 2011 annual conference on Human factors in computing systems*, 2011, p. 1031–1040.
- [7] Rukzio E, Holleis P, Gellersen H. Personal projectors for pervasive computing. *IEEE Pervasive Computing*, 2011.
- [8] Harrison C, Benko H, and Wilson AD. OmniTouch: wearable multitouch interaction everywhere. *In Proceedings of the 24th annual ACM symposium on User interface software and technology*, 2011, p. 441–450.
- [9] Mistry P, Maes P, Chang L. WUW-wear Ur world: a wearable gestural interface. *In Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, 2009, p. 4111–4116.
- [10] Spitzer MB, Rensing N, McClelland R, Aquilino P. Eyeglass-based systems for wearable computing. *In Wearable Computers, Digest of Papers, First International Symposium on*, 1997, p.48–51.
- [11] Kasai I, Tanijiri Y, Endo T, Ueda H. A forgettable near eye display. *In Wearable Computers, The Fourth International Symposium on*, 2000, p.115–118.
- [12] Tamaki E, Miyaki T, Rekimoto J. Brainy hand: an ear-worn hand gesture interaction device. *In Proceedings of the 27th international conference extended abstracts on Human factors in computing systems*, 2009, p. 4255–4260.
- [13] Ishii H. The tangible user interface and its evolution. *Communications of the ACM*;51: 6, 2008, p. 32–36.
- [14] Holman D, Vertegaal R, Altosaar M, Troje N, Johns D. Paper windows: interaction techniques for digital paper. *In Proceedings of the SIGCHI conference on Human factors in computing systems*, 2005, p. 591–599.
- [15] Mistry P, Maes P. Quickies: Intelligent sticky notes. *In Intelligent Environment, IET 4<sup>th</sup> International Conference*, 2008, p. 1–4.
- [16] Akaoka E, Ginn T, Vertegaal R. DisplayObjects: prototyping functional physical interfaces on 3d styrofoam, paper or cardboard models. *In Proceedings of the fourth international conference on Tangible, embedded, and embodied interaction*, 2010, p. 49–56.
- [17] Milgram P, Kishino F. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*;77:12, 1994, p.1321–1329.
- [18] Buxton W. *A three-state model of graphical input*. In: Diaper D, Gilmore DJ, Cockton G, Shackel B, editors. *In Proceedings of the IFIP Tc13 third international conference on human-computer interaction*, 1990. Amsterdam: North-Holland Publishing Co., p. 449–456.
- [19] Kato IPH, Billinghurst M, Poupyrev I. *Artoolkit user manual*. version 2.33. Washington : University of Washington, Human Interface Technology Lab, 2000.
- [20] Dul J, Weerdmeester BA. *Ergonomics for beginners: a quick reference guide*. CRC Press, 2008.
- [21] Allen IE, Seaman CA. Likert scales and data analyses. *Quality Progress*; 40: 7, 2007, p. 64–65.
- [22] Grant DA. The latin square principle in the design and analysis of psychological experiments. *Psychological bulletin*; 45: 5, 1948, p. 427.
- [23] Lehmann EL. *Nonparametrics: statistical methods based on ranks*. Prentice-Hall. 1st ed. 1975. Berlin: Springer. Revised ed. 2006.